

OIPE

## RAW SEQUENCE LISTING

DATE: 10/04/2001

PATENT APPLICATION: US/09/936,588

TIME: 14:36:58

Input Set : A:\163-pct.app

Output Set: N:\CRF3\10042001\I936588.raw

ENTERED

3 <110> APPLICANT: XU, Ming-Qun  
 4 EVANS, Thomas C.  
 5 PRADHAN, Sriharsa  
 6 COMB, Donald G.  
 7 PAULUS, Henry  
 8 SUN, Luo  
 9 CHEN, Lixin  
 10 GHOSH, Inca  
 11 NEW ENGLAND BIOLABS, INC.  
 12 BOSTON BIOMEDICAL RESEARCH INSTITUTE  
 14 <120> TITLE OF INVENTION: METHOD FOR GENERATING SPLIT, NON-TRANSFERABLE GENES  
 15 THAT ARE ABLE TO EXPRESS AN ACTIVE PROTEIN PRODUCT  
 17 <130> FILE REFERENCE: NEB-163-PCT  
 C--> 19 <140> CURRENT APPLICATION NUMBER: US/09/936,588  
 C--> 20 <141> CURRENT FILING DATE: 2001-09-13  
 22 <150> PRIOR APPLICATION NUMBER: 60/135,677  
 23 <151> PRIOR FILING DATE: 1999-05-24  
 25 <160> NUMBER OF SEQ ID NOS: 134  
 27 <170> SOFTWARE: PatentIn Ver. 2.0  
 29 <210> SEQ ID NO: 1  
 30 <211> LENGTH: 19  
 31 <212> TYPE: DNA  
 32 <213> ORGANISM: Escherichia coli  
 34 <400> SEQUENCE: 1  
 35 ggacggggaa ctaactatg 19  
 37 <210> SEQ ID NO: 2  
 38 <211> LENGTH: 20  
 39 <212> TYPE: DNA  
 40 <213> ORGANISM: Escherichia coli  
 42 <400> SEQUENCE: 2  
 43 ccacgatgac gcaccacgcg 20  
 45 <210> SEQ ID NO: 3  
 46 <211> LENGTH: 30  
 47 <212> TYPE: DNA  
 48 <213> ORGANISM: Escherichia coli  
 50 <400> SEQUENCE: 3  
 51 ggagggggca tatgaatggc gcacagtggg 30  
 53 <210> SEQ ID NO: 4  
 54 <211> LENGTH: 25  
 55 <212> TYPE: DNA  
 56 <213> ORGANISM: Escherichia coli  
 58 <400> SEQUENCE: 4  
 59 ggggggtcat gataatttct ccaac 25  
 61 <210> SEQ ID NO: 5  
 62 <211> LENGTH: 28  
 63 <212> TYPE: DNA  
 64 <213> ORGANISM: Escherichia coli

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/936,588

DATE: 10/04/2001

TIME: 14:36:58

Input Set : A:\163-pct.app

Output Set: N:\CRF3\10042001\I936588.raw

```

66 <400> SEQUENCE: 5
67 ccgggtggcg taattatgcc ggtttacg                28
69 <210> SEQ ID NO: 6
70 <211> LENGTH: 28
71 <212> TYPE: DNA
72 <213> ORGANISM: Escherichia coli
74 <400> SEQUENCE: 6
75 cgtaaaccgg cataattacg ccaccgcg                28
77 <210> SEQ ID NO: 7
78 <211> LENGTH: 14
79 <212> TYPE: PRT
80 <213> ORGANISM: Synechocystis PCC6803
82 <400> SEQUENCE: 7
83 Leu Glu Lys Phe Ala Glu Tyr Cys Phe Asn Lys Ser Thr Gly
84   1           5           10
87 <210> SEQ ID NO: 8
88 <211> LENGTH: 21
89 <212> TYPE: PRT
90 <213> ORGANISM: Escherichia coli
92 <400> SEQUENCE: 8
93 Cys Ala Gln Trp Val Val His Ala Leu Arg Ala Gln Gly Val Asn Thr
94   1           5           10           15
96 Val Phe Gly Tyr Gly
97           20
100 <210> SEQ ID NO: 9
101 <211> LENGTH: 20
102 <212> TYPE: PRT
103 <213> ORGANISM: Escherichia coli
105 <400> SEQUENCE: 9
106 Cys Val Trp Pro Leu Val Pro Pro Gly Ala Ser Asn Ser Glu Met Leu
107   1           5           10           15
109 Glu Lys Leu Ser
110           20
113 <210> SEQ ID NO: 10
114 <211> LENGTH: 26
115 <212> TYPE: DNA
116 <213> ORGANISM: Escherichia coli
118 <400> SEQUENCE: 10
119 ggggggtcatg aatggcgccac agtggg                26
121 <210> SEQ ID NO: 11
122 <211> LENGTH: 34
123 <212> TYPE: DNA
124 <213> ORGANISM: Escherichia coli
126 <400> SEQUENCE: 11
127 gcgcgctcga gttgatttaa cggctgctgt aatg                34
129 <210> SEQ ID NO: 12
130 <211> LENGTH: 32
131 <212> TYPE: DNA
132 <213> ORGANISM: Escherichia coli

```

## RAW SEQUENCE LISTING

DATE: 10/04/2001

PATENT APPLICATION: US/09/936,588

TIME: 14:36:58

Input Set : A:\163-pct.app

Output Set: N:\CRF3\10042001\I936588.raw

```

134 <400> SEQUENCE: 12
135 gcgcgaccgg ttgtgactgg cagcaacact gc 32
137 <210> SEQ ID NO: 13
138 <211> LENGTH: 31
139 <212> TYPE: DNA
140 <213> ORGANISM: Escherichia coli
142 <400> SEQUENCE: 13
143 ggggggctgc agtcatgata atttctccaa c 31
145 <210> SEQ ID NO: 14
146 <211> LENGTH: 22
147 <212> TYPE: DNA
148 <213> ORGANISM: MAIZE
150 <400> SEQUENCE: 14
151 atcagtacac agtcctgcc a tc 22
153 <210> SEQ ID NO: 15
154 <211> LENGTH: 20
155 <212> TYPE: DNA
156 <213> ORGANISM: MAIZE
158 <400> SEQUENCE: 15
159 gagacagccg ccgcaaccat 20
161 <210> SEQ ID NO: 16
162 <211> LENGTH: 29
163 <212> TYPE: DNA
164 <213> ORGANISM: MAIZE
166 <400> SEQUENCE: 16
167 gggcccatat ggccaccgcc gccgccgcg 29
169 <210> SEQ ID NO: 17
170 <211> LENGTH: 29
171 <212> TYPE: DNA
172 <213> ORGANISM: MAIZE
174 <400> SEQUENCE: 17
175 gggccctcga ggcttccttc aagaagagc 29
177 <210> SEQ ID NO: 18
178 <211> LENGTH: 29
179 <212> TYPE: DNA
180 <213> ORGANISM: MAIZE
182 <400> SEQUENCE: 18
183 gggccaccgg tacatcaaag aagagcttg 29
185 <210> SEQ ID NO: 19
186 <211> LENGTH: 31
187 <212> TYPE: DNA
188 <213> ORGANISM: MAIZE
190 <400> SEQUENCE: 19
191 ggggctgcat tcagtacaca gtcctgccat c 31
193 <210> SEQ ID NO: 20
194 <211> LENGTH: 7
195 <212> TYPE: PRT
196 <213> ORGANISM: Synechocystis PCC6803
198 <400> SEQUENCE: 20

```

## RAW SEQUENCE LISTING

DATE: 10/04/2001

PATENT APPLICATION: US/09/936,588

TIME: 14:36:58

Input Set : A:\163-pct.app

Output Set: N:\CRF3\10042001\I936588.raw

```

199 Leu Glu Lys Phe Ala Glu Tyr
200   1           5
203 <210> SEQ ID NO: 21
204 <211> LENGTH: 7
205 <212> TYPE: PRT
206 <213> ORGANISM: Synechocystis PCC6803
208 <400> SEQUENCE: 21
209 Cys Phe Asn Lys Ser Thr Gly
210   1           5
213 <210> SEQ ID NO: 22
214 <211> LENGTH: 21
215 <212> TYPE: PRT
216 <213> ORGANISM: MAIZE
218 <400> SEQUENCE: 22
219 Cys Lys Gly Ala Asp Ile Leu Val Glu Ser Leu Glu Arg Cys Gly Val
220   1           5           10           15
222 Arg Asp Val Phe Ala
223           20
226 <210> SEQ ID NO: 23
227 <211> LENGTH: 21
228 <212> TYPE: PRT
229 <213> ORGANISM: MAIZE
231 <400> SEQUENCE: 23
232 Cys Ile Pro Ser Gly Gly Ala Phe Lys Asp Met Ile Leu Asp Gly Asp
233   1           5           10           15
235 Gly Arg Thr Val Tyr
236           20
239 <210> SEQ ID NO: 24
240 <211> LENGTH: 44
241 <212> TYPE: DNA
242 <213> ORGANISM: Artificial Sequence
244 <220> FEATURE:
245 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
246     based on Salmonella typhimurium
248 <400> SEQUENCE: 24
249 ggatcctaag aaggagatat acccatggaa tccctgacgt taca           44
251 <210> SEQ ID NO: 25
252 <211> LENGTH: 38
253 <212> TYPE: DNA
254 <213> ORGANISM: Artificial Sequence
256 <220> FEATURE:
257 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
258     based on Salmonella typhimurium
260 <400> SEQUENCE: 25
261 gtcgacgctc tctgcagtt aggcaggcgt actcattc           38
263 <210> SEQ ID NO: 26
264 <211> LENGTH: 38
265 <212> TYPE: DNA
266 <213> ORGANISM: Artificial Sequence

```

## RAW SEQUENCE LISTING

DATE: 10/04/2001

PATENT APPLICATION: US/09/936,588

TIME: 14:36:58

Input Set : A:\163-pct.app

Output Set: N:\CRF3\10042001\I936588.raw

```

268 <220> FEATURE:
269 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
270     based on Salmonella typhimurium
272 <400> SEQUENCE: 26
273 gctttgctcc tggcggcttt acctgtggt aaaaccgc          38
275 <210> SEQ ID NO: 27
276 <211> LENGTH: 38
277 <212> TYPE: DNA
278 <213> ORGANISM: Artificial Sequence
280 <220> FEATURE:
281 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
282     based on Salmonella typhimurium
284 <400> SEQUENCE: 27
285 gcggttttac cacaaggtaa agccgccagg agcaaagc          38
287 <210> SEQ ID NO: 28
288 <211> LENGTH: 25
289 <212> TYPE: DNA
290 <213> ORGANISM: Artificial Sequence
292 <220> FEATURE:
293 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
294     based on Salmonella typhimurium
296 <400> SEQUENCE: 28
297 gccctaaag acacaattat tcgcg          25
299 <210> SEQ ID NO: 29
300 <211> LENGTH: 25
301 <212> TYPE: DNA
302 <213> ORGANISM: Artificial Sequence
304 <220> FEATURE:
305 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
306     based on Salmonella typhimurium
308 <400> SEQUENCE: 29
309 cagcggcgcc gtcacagca gagcg          25
311 <210> SEQ ID NO: 30
312 <211> LENGTH: 25
313 <212> TYPE: DNA
314 <213> ORGANISM: Artificial Sequence
316 <220> FEATURE:
317 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
318     based on Salmonella typhimurium
320 <400> SEQUENCE: 30
321 gcgaaccacc actaccaaca atttg          25
323 <210> SEQ ID NO: 31
324 <211> LENGTH: 25
325 <212> TYPE: DNA
326 <213> ORGANISM: Artificial Sequence
328 <220> FEATURE:
329 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic,
330     based on Salmonella typhimurium
332 <400> SEQUENCE: 31

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/936,588

DATE: 10/04/2001

TIME: 14:36:59

Input Set : A:\163-pct.app

Output Set: N:\CRF3\10042001\I936588.raw

L:19 M:270 C: Current Application Number differs, Replaced Application Number

L:20 M:271 C: Current Filing Date differs, Replaced Current Filing Date